



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 5, May 2023

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.423

 9940 572 462

 6381 907 438

 [ijrset@gmail.com](mailto:ijrset@gmail.com)

 [www.ijrset.com](http://www.ijrset.com)



# Efficacious Face Detection System based on Color and Depth of Skin

Ms.Poonam V. Kadam

Assistant Professor, Department of Computer Engineering, Metropolitan Institute of Technology & Management (MITM), Sindhudurg, Maharashtra, India

**ABSTRACT:**Face detection is the first step of any automated face recognition system. One of the most popular approaches to detect faces in color images is using a skin color & depth of skin, which in many cases needs a proper representation of color spaces to interpret image information. In this paper, we propose a fuzzy system for detecting skin in color images, so that each color tone is assumed to be a fuzzy set. The Red, Green, and Blue (RGB), the Hue, Saturation and Value (HSV), and the YCbCr (where Y is the luminance and Cb,Cr are the chroma components) color systems are used for the development of our fuzzy design. Thus, a fuzzy three-partition entropy approach is used to calculate all of the parameters needed for the fuzzy systems, and then, a face detection method is also developed to validate the segmentation results. The results of the experiments show a correct skin detection rate between 94% and 96% for our fuzzy segmentation methods, with a false positive rate of about 0.5% in all cases. Furthermore, the average correct face detection rate is above 93%, and even when working with heterogeneous backgrounds and different light conditions, it achieves almost 88% correct detections. Thus, our method leads to accurate face detection results with low false positive and false negative rates.

**KEYWORDS:** Adaboost Algorithm, Haar Cascade Classifier, OpenCV, Skin Color, Depth etc.

## I. INTRODUCTION

Over the last few years or so, face recognition has become a popular area of research in computer vision. Face recognition is also one of the most successful applications of image analysis and understanding. Because of the nature of the problem of face recognition, not only computer science researchers are interested in it, but neuroscientists and psychologists are also interested for the same. It is the general opinion that advances in computer vision research will provide useful insights to neuroscientists and psychologists into how human brain works, and vice versa.

The topic of real time face recognition for video and complex real-world environments has garnered tremendous attention for student to attend class daily means online attendance system as well as security system based on face recognition. Automated face recognition system is a big challenging problem and has gained much attention from last few decades. There are many approaches in this field. Many proposed algorithms are there to identify and recognize human being face form given dataset. The recent development in this field has facilitated us with fast processing capacity and high accuracy. The efforts are also going in the direction to include learning techniques in this complex computer vision technology. There are many existing systems to identify faces and recognized them. But the systems are not so efficient to have automated face detection, identification and recognition.

## II. ALGORITHMS

### ADABOOST ALGORITHM

“Boosting” is a general method for improving the performance of any learning algorithm. In theory, boosting can be used to significantly reduce the error of any “weak” learning algorithm that consistently generates classifiers which need only be a little bit better than random guessing. Despite the potential benefits of boosting promised by the theoretical results, the true practical value of boosting can only be assessed by testing the method on “real” learning problems. I present such an experimental assessment of a new boosting algorithm called **AdaBoost**.

Boosting works by repeatedly running a given weak learning algorithm on various distribution over the training data, and then combining the classifiers produced by the weak learner into a single composite classifier. The first provably effective boosting algorithms were presented by Schapire and Freund. More recently, I described and analyzed



**AdaBoost**, and we argued that this new boosting algorithm has certain properties which make it more practical and easier to implement than its predecessors.

**AdaBoost**, short for Adaptive Boosting, is a machine learning algorithm, formulated by Yoav Freund and Robert Schapire. It is a meta-algorithm, and can be used in conjunction with many other learning algorithms to improve their performance. AdaBoost is adaptive in the sense that subsequent classifiers built are tweaked in favor of those instances misclassified by previous classifiers[3]. AdaBoost is sensitive to noisy data and outliers. In some problems, however, it can be less susceptible to the overfitting problem than most learning algorithms. The classifiers it uses can be weak (i.e., display a substantial error rate), but as long as their performance is not random (resulting in an error rate of 0.5 for binary classification), they will improve the final model. Even classifiers with an error rate higher than would be expected from a random classifier will be useful, since they will have negative coefficients in the final linear combination of classifiers and hence behave like their inverses. AdaBoost generates and calls a new weak classifier in each of a series of rounds. For each call, a distribution of weights is updated that indicates the importance of examples in the data set for the classification. On each round, the weights of each incorrectly classified example are increased, and the weights of each correctly classified example are decreased, so the new classifier focuses on the examples which have so far eluded correct classification.

The basic idea of boosting is actually to use the weak learner of the features calculated, to form a highly correct prediction rules by calling the weak learner repeatedly processed on the different-different distributions over the training examples.

**HAAR CASCADE CLASSIFIER**

A Haar Classifier is also a machine learning algorithmic approach for the visual object detection, originally given by Viola & Jones. This technique was originally intended for the facial recognition but it can be used for any other object. The most important feature of the HaarClassifier is that, it quickly rejects regions that are highly unlikely to be contained in the object. The core basis for Haar cascade classifier object detection is the Haar-like features. These features, rather than using the intensity values of a pixel, use the change in contrast values between adjacent rectangular groups of pixels. The variance of contrast between the pixel groups are used to determine relative light and dark areas. The various Haar-like-features are shown in the figure 1. The set of basic Haar-like-feature is shown in figure 2, rotating which the other features can be generated. The value of a Haar-like feature is the difference between the sum of the pixel gray level values within the black and white rectangular regions, i.e.,

$$f(x) = \text{Sum black rectangle (pixel gray level)} - \text{Sum white rectangle (pixel gray level)}$$

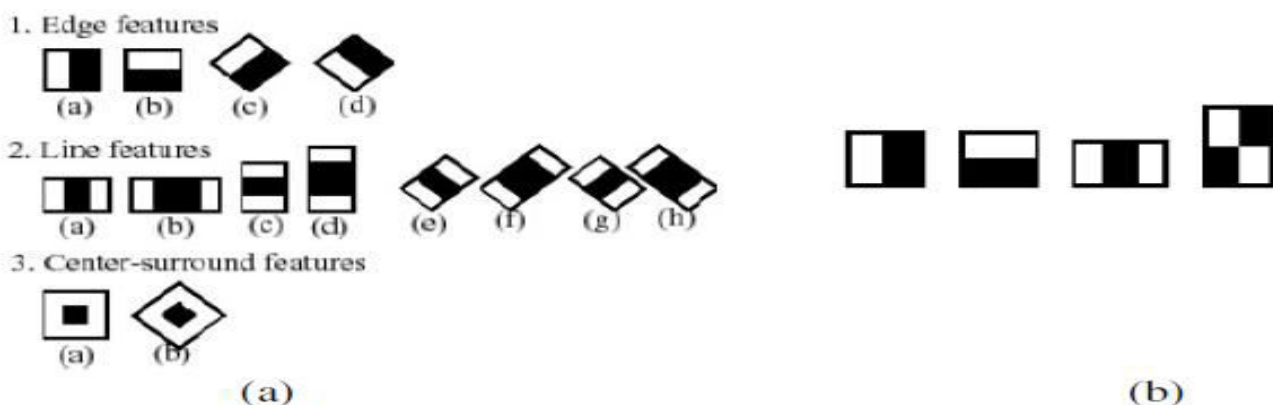


Fig 1. Haar-like-features

Comparing with the raw pixel values, Haar-like features can reduce/increase the in-class/out-of-class variability, and thus making classification much easier. The rectangle Haar-like features can be computed rapidly using “integral image”. Integral image at location of  $x, y$  contains the sum of the pixel values above and left of  $x, y$ , inclusive:

$$P(x, y) = \sum_{x' < x, y' < y} i(x', y')$$

$$P_1 = A, P_2 = A+B, P_3 = A+C, P_4 = A+B+C+D$$



|| Volume 12, Issue 5, May 2023 ||

| DOI:10.15680/IJIRSET.2023.1205173 |

$$P_1+P_4-P_2-P_3 = A+A+B+C+D-A-B-A-C=D$$

The sum of pixel values within “D”:

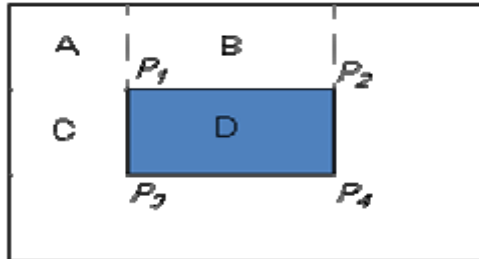


Fig2. Haar-like features

Using this Haar-like features the face detection cascade can be designed as in the figure below. In this Haar cascade classifier an image is classified as a human face if it passes all the conditions,  $\{f_1, f_2, \dots, f_n\}$ . If at any stage any of one or more conditions is false then the image does not contain the human face.

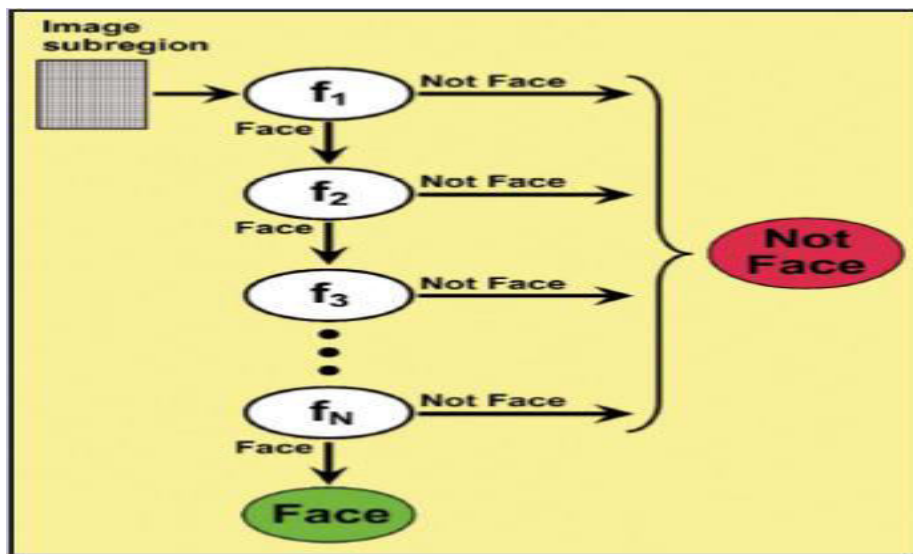


Fig3. The cascade classifier classified face and non-face.

The OpenCV is a great tool for image processing, and it provides several examples including face detection. In our work, for real-time process, we use OpenCV as a development tool to implement this algorithm.

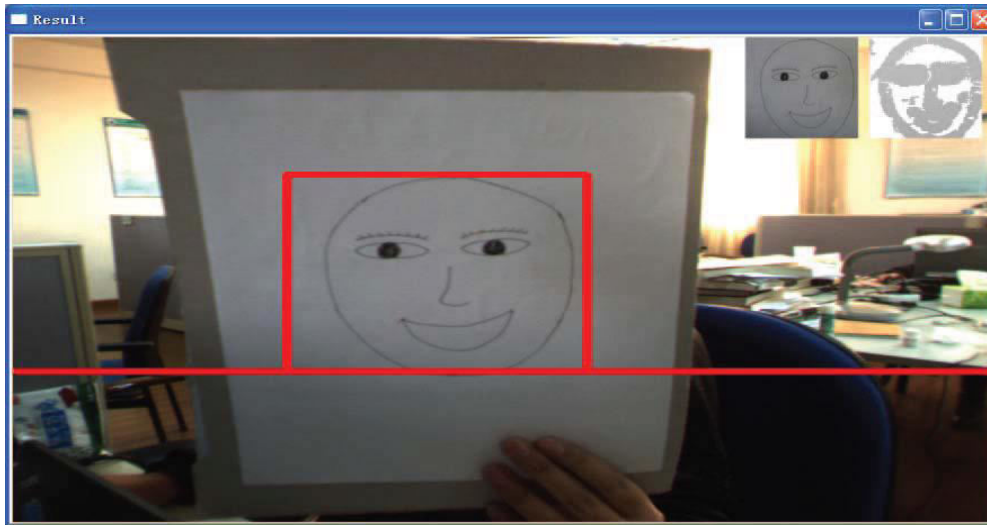


Figure 4. Example photo of face detection by using Adaboost algorithm

However, as shown in Fig.4, though the face area can be easily detected by using only Adaboost algorithm in OpenCV, it cannot distinguish fake face, such as picture face, from real human face. So, the detection algorithm needs to be improved for real-time applications such as vision guided service robot.

From the security point of view, the robot needs to distinguish a picture face and real one. It is really an important ability for the service robot to avoid being “cheated”. In order to avoid cheating by a picture face, skin color information should be considered. It will eliminate the problem of adaboost algorithm.

### III. METHODOLOGY

#### SKIN COLOR INFORMATION

**Human skin color** ranges in variety from the darkest brown to the lightest pinkishwhite hues. Human skin shows higher variation in color than any other single mammalian species and is the result of natural selection. Skin pigmentation in humans evolved to primarily regulate the amount of ultraviolet radiation penetrating the skin, controlling its biochemical effects.

The actual skin color of different humans is affected by many substances, although the single most important substance determining human skin color is the pigment melanin. Melanin is produced within the skin in cells called melanocytes and it is the main determinant of the skin color of darker-skinned humans. The skin color of people with light skin is determined mainly by the bluish-white connective tissue under the dermis and by the hemoglobin circulating in the veins of the dermis. The red color underlying the skin becomes more visible, especially in the face, when, as consequence of physical exercise or the stimulation of the nervous system (anger, fear), arterioles dilate.

The natural skin color can be darkened as a result of tanning due to exposure to sunlight. The leading theory is that skin color adapts to intense sunlight irradiation to provide partial protection against the ultraviolet fraction which produces damage and thus mutations in the DNA of the skin cells. There is a correlation between the geographic distribution of UV radiation (UVR) and the distribution of indigenous skin pigmentation around the world. Areas that highlight higher amounts of UVR reflect darker-skinned populations, generally located nearer towards the equator. Areas that are far from the tropics and closer to the poles have lower concentration of UVR, which is reflected in lighter-skinned populations. The example of variation in skin color is given below:



**What is Skin Color Detection?**

Skin detection is the process of finding skin-colored pixels and regions in an image or a video. This process is typically used as a preprocessing step to find regions that potentially have human faces and limbs in images. Several computer vision approaches have been developed for skindetection. A skin detector typically transforms a given pixel into an appropriate color space and then use a skin classifier to label the pixel whether it is a skin or a non-skin pixel. A skin classifier defines a decision boundary of the skin color class in the color space based on a training database of skin-colored pixels. The skin color mode in a color image can be given as:

$$r = R/(R+G+B)$$

$$g = G/(R+G+B)$$

$$Y = 0.30 R + 0.59G + 0.11 B$$

In which, the RGB values determine the original pixels. When values of *r*, *g*, *Y* satisfy the following formula, the color information in the skin area of human can be determined.

$$0.333 < r < 0.664$$

$$0.246 < g < 0.398$$

$$r > g$$

$$g \geq 0.5 - 0.5r$$

$$Y > 40$$

The advantage of this model is better color dot density detecting. By using this mode, the real faces with skin color information and the drawing faces without skin color information can be distinguished, as shown in Fig.2.3.1(a).



Fig5. Result of skin color detection

Skin detection in color images and videos is a very efficient way to locate skin-colored pixels, which might indicate the existence of human faces and hands. However, many objects in the real world have skin-tone colors, such as some kinds of leather, sand, wood, fur, etc., which might be mistakenly detected by a skin detector. Therefore, skin detection can be very useful in finding human faces and hands in controlled environments where the background is guaranteed not to contain skin-tone colors. Since skin detection depends on locating skin-colored pixels, its use is limited to color images, i.e., it is not useful with gray-scale, infrared, or other types of image modalities that do not contain color information. There is extensive research on finding human faces in images and videos using other cues such as finding local facial features or finding holistic facial templates. Skin detection can also be used as an efficient preprocessing filter to find potential skin regions in color images prior to applying more computationally expensive face or hand detectors. The skin color detection can eliminate general drawing faces without color of skin. But for real-face picture, it is hard to eliminate. To solve this problem, we use the 3D information of the face that is Depth Information.

### DEPTH INFORMATION

Face detection is one of the most thoroughly-explored questions in computer vision, with applications ranging from consumer cameras to human-computer interaction. It is also a vital skill for any robot designed to autonomously interact with humans. The faster and more reliable we can make face detection, the better our interaction with the people owning these faces will be. The computer vision community addresses face detection from a monocular image or video-centric perspective. Most algorithms are designed to detect faces using one or more camera images, without additional sensor information or context. A mobile robot, however, usually has multiple sensors in addition to its cameras, including sensors that can provide depth information. Laser range-finders, stereo camera pairs, the new Kinect sensor, and the Swiss Ranger camera can all provide depth information of varying types and qualities. Additionally, a mobile robot often knows the exact positions of its sensors and, hence, can calculate the 3D positions of the objects that it senses.

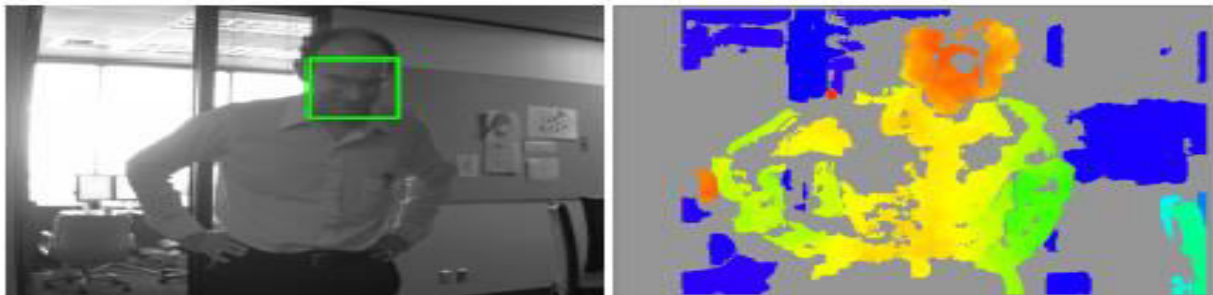


Fig 6 Image with face detection from (left); Corresponding stereo-generated depth image (right, false-color). Grey pixels lack depth information.

The Viola-Jones algorithm involves exhaustively searching an entire image for faces, with multiple scales explored at each pixel. Given no other knowledge about the scene, an exhaustive search is reasonable, and the Viola-Jones algorithm is efficient. If video data are available, the basic algorithm can be made more efficient by tracking the detected faces, and searching only a local neighborhood around faces found in previous frames. Extracting more contextual information, however, is difficult from monocular imagery.

According to the character that there is disparity field between two stereo images of the binocular camera, depth image of a face can be estimated. Fig. gives an example of the depth images and the right images obtained from of binocular camera for a specific person's face, what used here is the depth data.

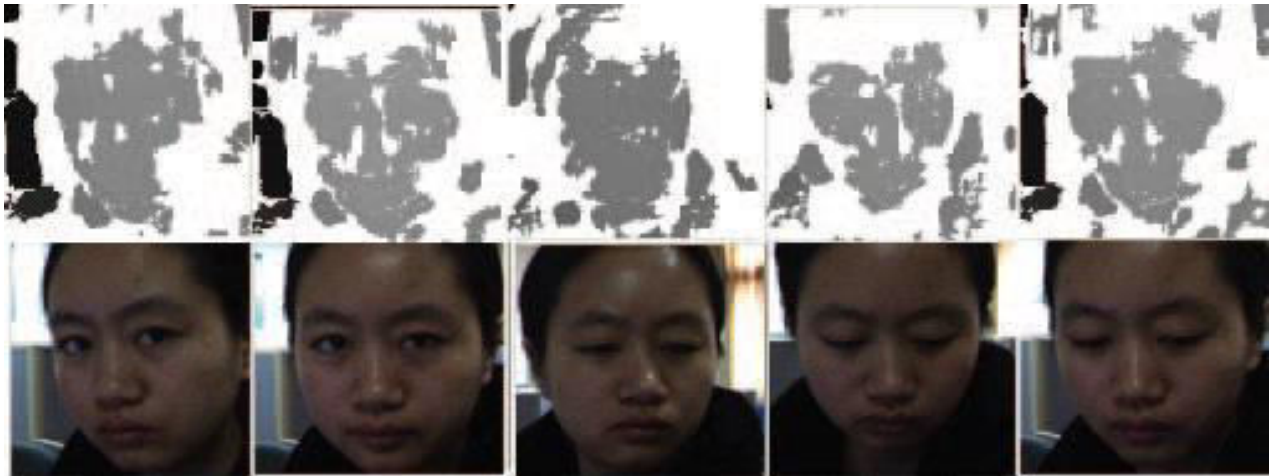


Fig 7. Depth images and right images of a face

Here we define  $D_i$  ( $i=1 \dots N$ ) as the depth value of the  $i$ th pixel in the face area, where  $N$  is the total number of pixels in the face area. As what can be observed in Figure, generally, the depth values of a real face has a much larger varied range than that of a picture face for plane of face. Considering this, we can define the average value of depth of all skin pixels,  $AvgD$ , as:

$$AvgD = \frac{\sum_{i=0}^N D_i}{N}$$

Then the variance of depth values of the face can be calculated as:

$$D_i = D_i - AvgD$$

$$S = \sqrt{\sum_{i=1}^N (d_i)^2}$$

In formula,  $d_i$  represents the difference between the depth value of each skin pixel and the average depth value,  $S$  gives the variance decision value of the detection result. Generally, the  $d_i$  varies much on a real face compared with a picture face for the same average value, as depicted in Figure. So, the value  $S$  is appropriate for decision.

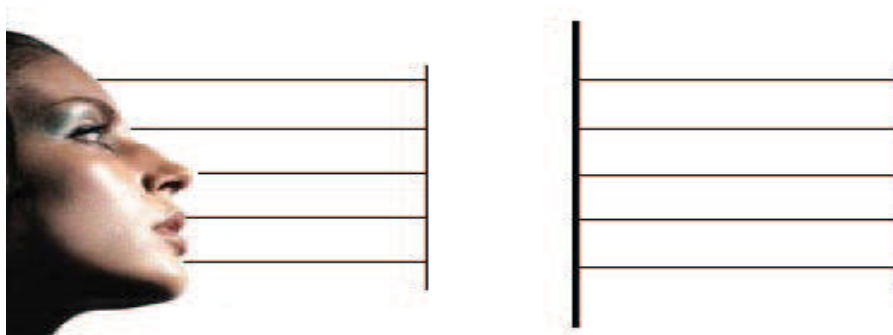


Fig8. Depth of real face (left) and picture face (right)



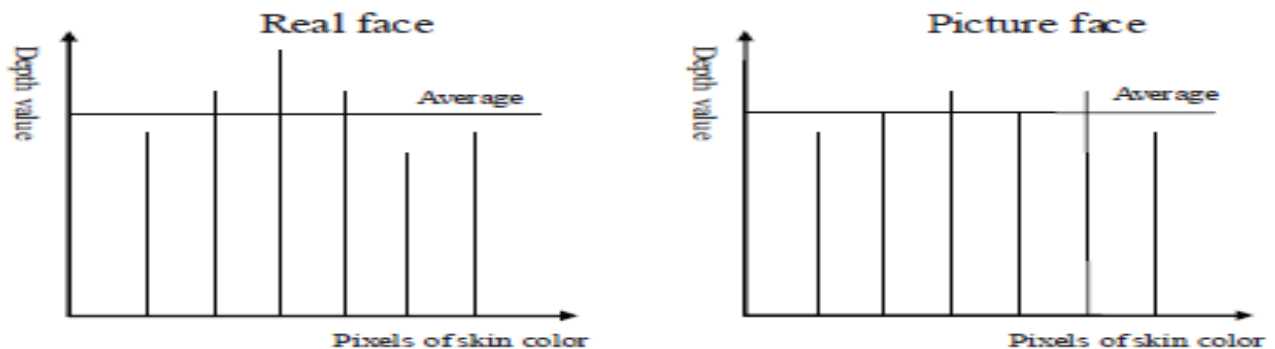


Fig 9. Distribution of the depth data of real face and picture face

#### IV. CONCLUSION

The proposed method improves 2D face detection techniques by additionally considering the facial 3D information obtained by binocular camera vision system. The skin color information and depth data of human face for detection and PCA algorithm for recognition are employed. It can not only detect person closed to the camera but also distinguish between real face and picture face. Applying the method to a service robot, a face which is faced to the camera can be determined whether it is a real face or a picture face, and then just do recognition on real face. The results show that this method is fast for the robot to do services, although the recognition rate is not high enough. Currently, the recognition accuracy may be affected by illumination, expressions and mechanical vibrations in some cases. This is left for future investigation.

#### REFERENCES

- [1] Brendan J. Frey, Antonio Colmenarez, and Thomas S.Huang, “Mixtures of local linear subspaces for face recognition,” in Proc. IEEEConference on ComputerVision and Pattern Recognition, 1998, pp. 32–37.
- [2] F.Tsalakanidou, D.Tzovaras, Use of depth and coloureigenfaces for face recognition. Pattern Recognition Letters 24, 2003, 1427–1435.
- [3] Li SZ, Zhu L, Zhang ZQ, Zhang HJ. “Learning to Detect Multi-View Faces in Real Time”. In Proeedings of the 2nd International Conference on Development and Leaming. Washington DC.2002.
- [4] P. Viola and M. Jones, Rapid object detection using a boostedcascade of simple features, in Proc. IEEE Computer Vision andPattern Recognition, 2001, p. 511.
- [5] T.-H. Sun, M. Chen, S. Lo, and F.-C. Tien. Face recognition using2D and disparity eigenface. Expert Syst.Appl.,vol.33,no.2, 2007,pp.265–273



SJIF Scientific Journal Impact Factor

Impact Factor: 8.423



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details